also act as a morphogen to control cell and tissue type specification. Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and is useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1738 of SEQ ID NO:30, b is an integer of 15 to 1752, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

This gene shares sequence homology with a yeast hypothetical 52.9 KD protein CDC26-YMR31 intergenic region (See Genbank Accession No. gplD50617lYSCCHRVI_114.). This gene has been mapped to chromosome 18q22-23, and therefore can be used in linkage analysis as a marker for 18q22-23.

This gene is expressed primarily in whole brain tissue, as well as brain specific tissues such as hypothalamus, frontal cortex, cerebellum, amygdala, and hippocampus tissues, as well as other brain specific tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to,schizophrenia, developmental disorders, and abnormal mental states. Similarly, polypeptides and antibodies directed to these polypeptidesare useful in providingimmunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, brain, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 131 as residues: Met-98 to Gln-107, Gly-120 to Gly-126, Pro-138 to Trp-145, Leu-159 to Gly-169, Val-211 to Arg-217, Cys-256 to His-262, Glu-320 to Val-327, Phe-399 to Asn-406, Asp-444 to Ser-450, Asp-475 to Trp-488.

The tissue distribution in whole brain tissue and brain specific tissues suggests that polynucleotides and polypeptides corresponding to this gene are useful for treating and/or diagnosing neural and neurodegenerative disorders. Furthermore, the tissue distribution suggests that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Elevated expression of this gene product within

the frontal cortex of the brain suggests that it may be involved in neuronal survival; synapse formation; conductance; neural differentiation, etc. Such involvement may impact many processes, such as learning and cognition. Additionally, the amygdala processes sensory information and relays this to other areas of the brain including the endocrine and autonomic domains of the hypothalamus and the brain stem. Thus, the translation product of this gene may also be useful for the detection and/or treatment of neural disorders that impact processes mediated by the amygdala. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2138 of SEQ ID NO:31, b is an integer of 15 to 2152, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

Preferred polypeptides of the invention comprise the following amino acid sequence: PPRPSTSGQWG (SEQ ID NO: 281) and/or RRSPFTSAQTG (SEQ ID NO: 282). Polynucleotides encoding these polypeptides are also provided.

The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in breast and soleus tissues, and, to a lesser extent, in several cell types, including T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are

225 230 235 240

Gln Val Gln Gly Pro Pro Lys Pro Arg Trp His Lys Asn Leu Thr Gly
245 250 255

Pro Gln Ile Ile Thr Leu Asn His Thr Asp Leu Val Pro Cys Leu Cys
260 265 270

Ile Gln Val Trp Pro Leu Glu Pro Asp Ser Val Arg Arg Thr Ser Ala
275 280 285

Pro Ser Gly Arg Thr Pro Ala His Thr Arg Thr Ser Gly Lys Pro Pro
290 295 300

Asp Cys Asp Cys Xaa
305

<210> 131
<211> 509
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (509)
<223> Xaa equals stop translation

<400> 131
Met Asp Pro Lys Leu Gly Arg Met Ala Ala Ser Leu Leu Ala Val Leu
1 5 10 15

Leu Leu Leu Leu Leu Glu Arg Gly Met Phe Ser Ser Pro Ser Pro Pro
20 25 30

Pro Ala Leu Leu Glu Lys Val Phe Gln Tyr Ile Asp Leu His Gln Asp
35 40 45

Glu Phe Val Gln Thr Leu Lys Glu Trp Val Ala Ile Glu Ser Asp Ser
50 55 60

Val Gln Pro Val Pro Arg Phe Arg Gln Glu Leu Phe Arg Met Met Ala
65 70 75 80

Val Ala Ala Asp Thr Leu Gln Arg Leu Gly Ala Arg Val Ala Ser Val
85 90 95

Asp Met Gly Pro Gln Gln Leu Pro Asp Gly Gln Ser Leu Pro Ile Pro
100 105 110

Pro Val Ile Leu Ala Glu Leu Gly Ser Asp Pro Thr Lys Gly Thr Val
115 120 125

Cys Phe Tyr Gly His Leu Asp Val Gln Pro Ala Asp Arg Gly Asp Gly
130 135 140

Trp Leu Thr Asp Pro Tyr Val Leu Thr Glu Val Asp Gly Lys Leu Tyr
145 150 155 160

Gly Arg Gly Ala Thr Asp Asn Lys Gly Pro Val Leu Ala Trp Ile Asn
165 170 175

Ala Val Ser Ala Phe Arg Ala Leu Glu Gln Asp Leu Pro Val Asn Ile
180 185 190

```
Lys Phe Ile Ile Glu Gly Met Glu Glu Ala Gly Ser Val Ala Leu Glu
        195                 200                 205

Glu Leu Val Glu Lys Glu Lys Asp Arg Phe Phe Ser Gly Val Asp Tyr
    210                 215                 220

Ile Val Ile Ser Asp Asn Leu Trp Ile Ser Gln Arg Lys Pro Ala Ile
225                 230                 235                 240

Thr Tyr Gly Thr Arg Gly Asn Ser Tyr Phe Met Val Glu Val Lys Cys
                245                 250                 255

Arg Asp Gln Asp Phe His Ser Gly Thr Phe Gly Gly Ile Leu His Glu
            260                 265                 270

Pro Met Ala Asp Leu Val Ala Leu Leu Gly Ser Leu Val Asp Ser Ser
        275                 280                 285

Gly His Ile Leu Val Pro Gly Ile Tyr Asp Glu Val Val Pro Leu Thr
    290                 295                 300

Glu Glu Glu Ile Asn Thr Tyr Lys Ala Ile His Leu Asp Leu Glu Glu
305                 310                 315                 320

Tyr Arg Asn Ser Ser Arg Val Glu Lys Phe Leu Phe Asp Thr Lys Glu
                325                 330                 335

Glu Ile Leu Met His Leu Trp Arg Tyr Pro Ser Leu Ser Ile His Gly
            340                 345                 350

Ile Glu Gly Ala Phe Asp Glu Pro Gly Thr Lys Thr Val Ile Pro Gly
        355                 360                 365

Arg Val Ile Gly Lys Phe Ser Ile Arg Leu Val Pro His Met Asn Val
    370                 375                 380

Ser Ala Val Glu Lys Gln Val Thr Arg His Leu Glu Asp Val Phe Ser
385                 390                 395                 400

Lys Arg Asn Ser Ser Asn Lys Met Val Val Ser Met Thr Leu Gly Leu
                405                 410                 415

His Pro Trp Ile Ala Asn Ile Asp Asp Thr Gln Tyr Leu Ala Ala Lys
            420                 425                 430

Arg Ala Ile Arg Thr Val Phe Gly Thr Glu Pro Asp Met Ile Arg Asp
        435                 440                 445

Gly Ser Thr Ile Pro Ile Ala Lys Met Phe Gln Glu Ile Val His Lys
    450                 455                 460

Ser Val Val Leu Ile Pro Leu Gly Ala Val Asp Asp Gly Glu His Ser
465                 470                 475                 480

Gln Asn Glu Lys Ile Asn Arg Trp Asn Tyr Ile Glu Gly Thr Lys Leu
                485                 490                 495

Phe Ala Ala Phe Phe Leu Glu Met Ala Gln Leu His Xaa
            500                 505
```

<210> 132
<211> 507

invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these

fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the